DOCUMENT RESUME

ED 480 063 CG 032 636

AUTHOR	Rubenstein, Jeff
TITLE	Test Preparation: What Makes It Effective?
PUB DATE	2003-08-00
NOTE	21p.; In: Measuring Up: Assessment Issues for Teachers, Counselors, and Administrators; see CG 032 608.
PUB TYPE	Information Analyses (070)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	College Admission; *College Entrance Examinations; *Educational Testing; *High Stakes Tests; *Program Effectiveness; *Test Coaching; Test Wiseness

ABSTRACT

As testing has taken on a more significant role in the college admission process and, more recently, in K-12 education through the increasing popularity of high-stakes testing, so have the urgency of questions surrounding test preparation initiatives, through both private and educational ventures and school-sponsored preparation programs. This chapter reviews the debate concerning test preparation program effectiveness and presents components necessary to a good test preparation program. (GCP)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 480 063

Test Preparation: What Makes It Effective?

By
Jeff Rubenstein

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

CG032636

Chapter 28
Test Preparation
What Makes It Effective?

Jeff Rubenstein

As testing has taken on a more significant role in the college admissions process and, more recently, in K–12 education through the increasing popularity of high-stakes testing, so have the urgency of questions surrounding test preparation initiatives, through both private educational ventures and school-sponsored preparation programs. As their popularity increases, no small amount of ink has been spilled and controversy raised about such programs—whether they are fair or not, whether they represent genuine educational enrichment or are somehow "gaming" the system, and whether they are ultimately effective.

Despite the controversy, the debate over the effectiveness of such programs is not a very interesting one. The reason for this is that the question of effectiveness has rarely been posed in an informative way. On the one hand, everyone (even those who are opposed to independent test preparation companies) agrees that test preparation, construed in the broadest sense, works. Where people disagree is on the details: whether a particular course of instruction is effective and exactly what about it is or is not effective. This is to say that the interesting question is not *whether* test preparation works, but *why* and *how* it does: Which aspects in an educational process (whether in school or in a private setting) prepare students adequately for their tests, and how can these best be implemented? This is the intelligent way to pose a question that is worth investigating.

I should make three introductory points about the goals of this chapter before proceeding: First, it is not my intention in this article to propound the virtues of one particular program of test preparation over another; I am writing with as much objectivity as possible given my necessarily interested position in this debate as an employee of Princeton Review. This means that I am keeping the discussion at a relatively general level, without going into specifics of particular programs. Even at this general level, however, there are certain overarching principles of good test preparation that the data show to be effective, and these

principles are worth elucidating. Second, for purposes of this article, I use SAT preparation as an example, because this is arguably the most prevalent and best known form of test preparation. In general the principles that apply to SAT preparation also apply, with some modification, to preparation for other standardized tests. Finally, as this anthology is intended for a general audience, I am keeping the discussion as nontechnical as possible, with minimal references. A short guide to further reading can be found at the end of the chapter.

Brief History of the Debate over the Effectiveness of SAT Preparation

The history of this debate can without too much injustice be characterized as one in which the Educational Testing Service (ETS) and the College Board have tried in various ways to discount an obvious fact: that test preparation is, in many cases, extremely effective. Oddly enough, even these critics of test preparation seem to agree that relatively short-term preparation can be quite useful. In fact, ETS internal research has shown that such preparation can be highly effective, and for years the College Board itself proudly displayed a quotation from a student who claimed an increase of 200 points through the use of its products. Nonetheless, these organizations' official position has always been that the gains are small, though somewhat larger for more robust programs (in the range of 40 hours).

When the original studies showing that relatively short-term test preparation could be effective in raising SAT scores were published in the 1970s, ETS and the College Board naturally perceived these findings as a threat to the legitimacy of the SAT. After all, if a few months of preparation could significantly influence scores, then how much "intelligence" could the test be measuring? ETS and the College Board initially responded to such studies by fighting a legal battle against the companies that offered test preparation services. They asked the Federal Trade Commission (FTC) to investigate these companies, accusing them of making false and misleading claims of score improvements.

The ensuing FTC investigation not only failed to indict these programs for fraudulent advertising, but actually provided evidence that these programs were effective: Their report showed that these programs, on average, raised scores by more than 100 points on the SAT combined (Levine, 1979). The FTC went on to state that the ETS was making fraudulent claims in asserting that test preparation was not effective. In fact, Albert Kramer, director of the Bureau of Consumer

Protection, stated that in direct contradiction to what ETS and the College Board were claiming, coaching could be efficacious.

ETS researchers have also published their own studies showing that test preparation is effective. In 1978, one such researcher, Lewis Pike, made a comprehensive study of all available data, both internal and external, on the subject. He reported that the SAT Math section was highly coachable and that the SAT Verbal was also likely so (but the evidence was still insufficient to show this definitively). Shortly after issuing his report, Pike departed from ETS.

The more recent ETS and College Board strategy to discount the effectiveness of test preparation programs has been somewhat more subtle. Recently, they have begun to admit the existence of score improvements while attempting to deny that these improvements were the result of test preparation narrowly defined (which they refer to as "coaching"). This strategy involves two interesting moves that make their studies, although not statistically wrong, certainly misleading. The first move is to throw all kinds of test preparation programs into one category and study them together as if they were comparable. When ETS now does a study on "test preparation," it defines the construct so broadly as to encompass well-researched, 40-hour programs and poorly designed, one- or two-hour after-school preparation sessions. As a result, ETS studies conclude that test preparation does not work. In a way no conclusion could be more obvious; an hour or two of unprofessional preparation is almost certain to be of little use. But the question posed by the study is not very interesting: No useful study can be conducted with constructs that are so broad as to encompass both well-tested, rigorous programs of test preparation and one- or two-hour presentations that call themselves test preparation. If the effectiveness of test preparation is to be studied seriously, such a study will have to be carried out on a specific, well-defined form of test preparation, with rigorous standards as to what this sort of preparation constitutes and what it does not.

The second move made in the more recent studies—after lumping all kinds of test preparation programs together—is to try to separate the parts of any test preparation program that ETS considers "good" preparation (familiarity with question types, general guessing strategies, etc.) from those they consider "bad" preparation (what ETS calls coaching; namely, instruction in test-specific problem-solving skills).

There is more than a little sleight of hand here in the way ETS slices test preparation into "coaching," "preparation," and "enrichment." By removing elements that it now claims are normal and rational kinds

of preparatory activities (though the idea that any sort of preparation is a good idea is a very recent discovery on its part), ETS tries to adjust the improvement numbers statistically so as to reduce the perceived net effectiveness of a test preparation program. That is, more recent studies try to chip away at the whole (rather substantial) improvement number by statistically removing the effect of several important elements (such as any genuine educational enrichment and the beneficial effect of timed practice tests). Such studies are misleading because any good preparation program involves many elements, each of which contributes to the effectiveness of the whole program. They are misleading also because ETS has effectively defined test preparation as something that will have no effect: It is committed to the proposition that coaching is ineffective, though genuine enrichment may indeed raise scores. Therefore, if any program does succeed in improving scores, then what that program must be doing is not coaching but rather genuine enrichment. This amazing piece of circular reasoning is what underlies ETS's current view of test preparation programs and its attempt to reduce how effective these programs appear on paper by reclassifying much of the content of such programs under the rubrics of "enrichment" or "preparation."

Given that none of these recent studies—and no proposed study that I am aware of—is going to tell us much of anything interesting about test preparation in general, what can we say of interest about it? In such circumstances, one rational approach is to begin with the evidence closest to hand, then begin to investigate its possible implications. That is, we should start with the following fact: In some cases test preparation, broadly construed, clearly does work well—sometimes extraordinarily so. In other cases, it is less effective. And many cases lie somewhere between the two extremes on this continuum of effectiveness. The proponents and detractors of test preparation all agree on this point (though they may disagree on how much of the effect is due to enrichment or coaching or something else).

The interesting question then becomes, In the cases where test preparation works, *why* does it work? What concrete problems are being addressed by effective test preparation, and how can we systematically address them in the future such that all students have a fair testing experience and are in a position to perform to the best of their abilities?

Areas Where Test Preparation Can Improve Performance

One useful way to begin discussing the reasons why test

preparation is effective is to discuss the problems that unprepared students have when taking standardized tests. There are a great number of such problems, which for our purposes I group into five categories: (a) poor general testing strategies (pacing, question selection, and setting priorities); (b) lack of specific problem-solving skills, which relates to the fact that particular questions on a test instrument are not aligned with a student's learned curriculum; (c) lack of practice with the preceding skills and lack of ability to deal with the psychological difficulties attached to a standardized test; (d) physical exhaustion; and (e) lack of basic skills that were part of a student's learned curriculum. Let us take a closer look at each of these in turn.

Poor Testing Strategies

ETS researcher Franklin Evans (1980) noted in a study that many of his test students, especially minority students, were probably underperforming on his SAT-like tests because they were using inefficient test-taking strategies. What exactly do such strategies constitute?

To formulate a solid test-taking strategy for any particular test requires, at a minimum, (a) understanding how the test is scored, (b) having a timing strategy, and (c) having a question selection and priority setting strategy. Each of these needs to be tailored to the particular characteristics of the test being taken.

Understanding How a Test Is Scored

The particular way a test is scored has a significant influence on the most rational way to approach various aspects of testing. Certain tests have deductions for incorrect answers, which are intended to neutralize the effect of random guessing; other tests do not have such deductions. This distinction makes a large difference when it comes to formulating the most intelligent guessing strategy for a given test instrument. Other tests have essay portions (which may or may not have a specific time frame dedicated to them); students need to know the relative importance of such essays, and how much time they should devote to them relative to the multiple-choice parts of the test.

Further, the newer computer-adapted tests are very particular in their method of scoring and require a much more finely tuned sense of how to achieve the best score. Knowing when it is rational to guess, which parts of the test are worth more or fewer points, and where spending one's time yields the greatest benefit are crucial strategies

that are rarely explained in a clear way by the testing companies. A lack of understanding of how a test is scored leads to less-than-optimal pacing and guessing strategies, which may cost a student valuable points.

Timing

Most standardized tests are highly speeded; that is, only a small percentage of test takers are supposed to be able to finish them. How one deals with this fact can make an enormous difference in one's score. Given the short time allowed per question, it is impossible for many students to complete every question on a test. Many students will therefore opt for one of two simple (but often damaging) strategies:

1. Try to rush through the test, responding to every question, even though he or she doesn't have enough time to answer any of them thoroughly. This student spreads himself or herself too thin and may suffer a significant score reduction in consequence.
2. Go through the test at a more moderate pace, starting with the first question and working in order until time runs out. This student might finish questions 1 through 15 or 1 through 20 or 1 through 25, never knowing if there were easier questions later in the test that he or she never even attempted. This may also lead to a significant reduction in score.

Students need to learn how to cope with the highly speeded test environment, how to set priorities in problem solving in order to choose the questions that are most advantageous for them to answer, and how to spend the appropriate amount of time on each question. They also need to learn how best to do this given the particular structure of particular tests: A very different strategy will be warranted depending on how any given test instrument is constructed.

Selecting Questions and Setting Priorities

Those students (the majority) for whom answering every question is counterproductive need a strategy for choosing which questions to answer, and in which order, to make the most beneficial use of their time. Students should usually choose the easier questions to answer first—giving them a better chance of spending their time on problems they will answer correctly—and save the more difficult ones for last (or not answer them at all). Certain tests are constructed with easier

questions in identifiable places; other tests do not have an identifiable order of difficulty, which means that students need other tools to help them select and prioritize questions. In either case, part of any good testing strategy is knowing which questions to answer, in what order, and how much time is worth spending on each of them; and a lack of understanding of these points puts students at a significant disadvantage.

Problem-Solving Skills Needed for Questions Not Covered in the Curriculum

Probably the most significant problem with many current tests is the lack of alignment with typical school curricula. That is to say, these tests are in some fashion testing something other than what students are being taught. This is usually not so much a matter of conceptual differences as of application. A particular instrument may test the same basic rules of geometry that students learn in school but test them in a fundamentally different way than students have learned them. This is another kind of misalignment with curriculum, one that has a significant effect on students' ability to show what they know.

A *good* test is a very important tool in the educational process. It helps to inform both teachers and students about how they are doing and gives direction as to how to improve. That is, it gives useful information to answer the question, For any given student, how much has that student learned relative to what should have been learned? The test thereby also indicates what the student has left to learn, and tells student and teacher what particular areas to focus on for improvement.

However, a test can be useful in this way only if the educational objectives are clearly outlined in advance, if the curriculum is well designed around those objectives, and if the test is aligned with that curriculum. Sadly, many tests, including the SAT, are not even remotely curriculum aligned (the SAT, for instance, shows its origins in the world of IQ tests on its face), and this creates a whole series of problems.

To the degree a test tests something other than what is covered in a student's learned curriculum (either by testing concepts that were not covered or by testing them in ways that were not covered) students require some additional preparation to perform up to their abilities. There are three ways in which the tested curriculum effectively goes beyond the learned curriculum, creating difficulties for students who are now being tested on matters for which they might not have been prepared: (a) employing particular question types that have not been covered in a school curriculum, (b) testing concepts that have not been covered in a school curriculum, and (c) testing known concepts in

fundamentally different ways than they have ever been presented in a school curriculum.

A good example of the first type is the Quantitative Comparison question type on the SAT. A Quantitative Comparison question looks like this:

Column A	*Column B*
x	x^2

A student is shown two columns of information (occasionally with additional information provided in the center) and is asked to pick a choice based on which of the following obtains:

1. Column A is greater than Column B.
2. Column B is greater than Column A.
3. The two columns are equal.
4. The relationship cannot be determined based on the information given.

For students who have not seen a problem of this type before (and even for those who have), it is not at all obvious what they are being asked to do. In particular, what does choice (4) "The relationship cannot be determined" mean? (In the preceding problem, the answer is in fact [4]. If the variable x is 0, then the two columns are each equal to 0; however, if x is equal to 2, then the value in column B is greater. Therefore the relationship cannot be determined.)

But, more importantly, even after these directions are explained to a student, is the student any better equipped to solve this sort of problem? Systematic instruction in this sort of problem solving is not part of any school curriculum of which I am aware (and it probably should not be). But if we are to ask students to perform well on this sort of test item, they should be given some sort of systematic instruction in solving it.

Let us move on to a problem of the second type: the testing of known concepts in very strange (and occasionally counterintuitive) ways. Anyone who has looked at an SAT carefully will be struck by the apparent simplicity of the concepts it tests. It is in fact true that the SAT Math section tests concepts from seventh-, eighth-, and ninth-grade mathematics only. How is it possible, then, that some problems on the SAT could seem so difficult? Not because the concepts are difficult, but because the problems present very simple concepts in very odd

ways. Here is an example of such a problem, taken from an actual SAT form:

Column A		*Column B*
	$3x + y = 15$ $x + 2y = 16$	
$2x - y$		0

This question effectively asks students to establish the value of $2x - y$ given two equations, each of which contains the variables x and y, then to determine whether that value is greater than, equal to, or less than 0. What is curious about this problem (aside from the Quantitative Comparison format) is that it actually works against students who "follow the rules." How? On its face, this question is formatted to look like a simultaneous equation problem. Most students are taught to solve simultaneous equation problems in the following manner: Multiply or divide one equation by a certain factor in order to allow one variable to be removed; solve for a single variable, then solve for the other. That is, do the following:

$$\begin{aligned} 3x + \ \ y &= 15 \\ x + 2y &= 16 \end{aligned}$$

In order to make the y variable drop out, we would multiply the first equation by 2, which gives us $6x + 2y = 30$. Now we can subtract one equation from the other:

$$\begin{aligned} 6x + 2y &= 30 \\ -x + 2y &= 16 \\ 5x \quad &= 14 \end{aligned}$$

This allows us to solve for x, which yields $\frac{14}{5}$. Now we can use this value for x in either equation to solve for y. Sadly, this is a very messy affair and a very long way to figure out the value of $2x - y$. There is, of course, a shortcut: simply to subtract one equation from the other:

$$\begin{aligned} 3x + \ \ y &= 15 \\ -x + 2y &= 16 \\ 2x - \ \ y &= -1 \end{aligned}$$

Many students, however, will never see this shortcut, primarily because the problem is designed so they will *not* see it if they follow the traditional rules for solving simultaneous equations. That is, students who assume that they should apply the rules they learned in math class will likely do far worse on this problem (and others like it) than students who do not.

This is a perfect illustration of how a student who may be fully competent at a basic skill—such as solving simultaneous equations—will still underperform on the SAT because he or she (even while knowing full well how to solve simultaneous equations) does not understand the particular quirky way that simultaneous equation problems are tested on the SAT.

Now let us turn to the third instance, in which a test actually tests concepts that are not covered in any school curriculum. Here are two more examples, both taken from the SAT:

> If a rectangular piece of paper is cut into exactly three pieces by making two straight cuts, which of the following could be the total number of edges on the three pieces?
> I. 9
> II. 11
> III. 12
>
> 1. I only
> 2. III only
> 3. I and III only
> 4. II and III only
> 5. I, II, and III

What skill exactly is this question supposed to be testing? Some strange spatial intuition? If one tries to stretch the idea, a case might be made that this falls under "basic geometric intuition or understanding" but that would largely be an after-the-fact rationalization. What this question really tests is whether a student can follow complex directions, experiment with various ways of making two straight cuts in a rectangular piece, and count up the resulting edges.

If you are not yet convinced that the SAT contains questions that are not curriculum aligned, take a look at this final example:

> 5:05
>
> The 12-hour digital clock above shows one example of a time at which the number representing the hour is equal to the

number representing the minutes. What is the *least* possible number of minutes from the instant one such "double" reading appears to the instant the next appears?

1. 11
2. 30
3. 49
4. 60
5. 61

Psychological Factors

Although the effect is difficult to quantify, one more factor in poor student performance is undoubtedly psychological. There is an enormous intimidation factor at work in these tests; the more high-stakes the test, the more pressure students feel, and in some cases, the worse they will perform. Add to that the fact that many of these tests are not curriculum aligned, and the result is that students are led to harmful thoughts such as, "This test supposedly only measures seventh-, eighth-, and ninth-grade math; if I can't solve these problems, I must be really stupid." This leads to the perception (which is occasionally correct) that the test questions rely on some bizarre intuition—either students "get it" or they don't—and that there is no relationship between how hard they have worked in school and their performance on the test. Many students who are perfectly adept math students in school but whose testing skills are weak are humiliated by such tests, largely because the tests are not well aligned with their school curriculum.

Most students, in fact, feel that the lack of testing skills that manifests itself in a low test score is an indication of their own lack of ability or a deficiency in themselves as students. The reality may be quite different: Although in certain cases students have simply not learned their lessons, quite often the problem lies in the fact that a question item is not aligned with their school curriculum or that they have never learned the testing strategies required to perform well on a standardized test. This lack of understanding about where their difficulty lies (i.e., that it lies in poor testing skills, not in poor mathematical or linguistic skills) adds to the frustration that many students feel with such tests and to their sense that such tests are unfair. This leads in some cases to a sense of helplessness and a difficulty in preparing seriously to perform well on such instruments.

In the case of minority students, even more significant problems may be at work. Claude Steele, a psychologist at Stanford, has done

some interesting work that makes the case that minority test takers, when told that minorities do poorly on a particular test, in fact do worse on that test than when they are told that minorities do equally well as other students on the test. That is, knowing that a test may be biased in fact produces subnormal results for certain groups of people.

Physical Exhaustion

Another reason why students may do poorly on standardized tests comes down to a simple factor of physical and mental exhaustion. Many students are wholly unprepared for the intensity of a testing experience, which demands a great deal of concentration, consistently applied, for quite a long time.

Basic Skills Problems

Finally, students may do poorly on standardized tests because of a genuine lack of the skills that the test is supposed to measure, such as basic mathematical skills. Of course, we have seen that in reality most tests are blunt instruments that measure many factors having nothing to do with basic skills. In many cases, however, a tester will perform poorly due to a genuine lack of certain basic skills.

The Solutions: What Goes Into a Good Test Preparation Program?

If the preceding list (although probably not complete) represents a good number of the factors that cause students to test poorly, then a good test preparation program would be one that addresses a substantial number of these problems, and an ideal program would be one that addresses all of them. That is, at least the following five elements are necessary to a thorough test preparation program:

1. Review of basic skills
2. Instruction in test-specific problem-solving skills, to give students the problem-solving skills they need to tackle non-curriculum-aligned test items
3. Instruction in overall test strategies, to ensure they have a coherent overall strategy (including pacing, question selection, and priority setting)
4. Delivery of full-length practice tests, to allow students to practice the skills learned in (2) and (3), as well as to increase their stamina and desensitize them to the stress

of a testing environment

5. An educational environment specifically targeted toward improving test-taking skills, to show students they are learning a specific task that they have not learned before (and that the lack of these skills is not their fault)

Of course, not every student will need all these elements; however, the more complete the program, the better chance it will have of promoting greater improvement in a greater number of students. Let us briefly consider each element in turn.

Review of Basic Skills

First, there should be a thorough review of those basic skills to be tested on a particular instrument, such as basic mathematical concepts, rules of English grammar, and other skills appropriate to the test in question. Such a review ensures that any gaps or forgotten bits of information get recapitulated in the weeks before testing.

Many students (especially those with well-designed school curricula and well-executed instruction) will gain something but not a tremendous amount from this aspect of the preparation, because they will already have adequate command of all the basic skills tested. The aspects that follow, those that deal with strategies and test-specific skills, are generally the more important factors in a program of test preparation.

Test-Specific Problem-Solving Skills

As we have seen in the preceding example problems from the SAT, any time a particular testing instrument diverges from the school curriculum (i.e., the tested curriculum is not aligned with the taught curriculum) a problem exists. Not only may certain skills be tested that were never taught in school (though these skills ideally would be addressed with supplemental work during the review of basic skills) but, more importantly, basic skills that students have mastered may be tested in ways that students have not mastered. That is, a student may be perfectly adept at a particular skill (e.g., solving simultaneous equations) but have no approach for putting that skill into practice in the particular way that skill is tested.

This mismatch occurs because, as we have seen in the preceding SAT examples, a test may present questions with a very particular sort of logic to them, a logic that students have never been taught and for which they need to learn test-specific problem-solving skills. This means that they need to learn how to put their knowledge into practice given

the way that questions are written on that particular instrument.

This is, of course, the reason why test preparation companies (due in part no doubt to their own rhetoric) are accused of merely teaching "tricks"—an accusation that I think is entirely false. Testing companies do indeed teach students how to solve problems in the particular way that they show up on an individual test. These skills are exactly what most students have never learned, and need to learn, to be able to perform to the best of their abilities on such tests. These skills are test-specific (and have little application outside the test itself), but this is not a result of the test preparation companies trying to "game" the system; it is a result of a test that asks questions in a peculiar and nonstandard fashion and therefore requires students to approach problems in a quirky fashion. For instance, a test preparation program might well teach students that when faced with a simultaneous equation problem, they should check to see whether a simple addition or subtraction operation will get them the desired result, instead of assuming that the question should be solved in the traditional way. (More often than not, simultaneous equation problems are structured this way on the SAT.) Students need to learn this test-specific skill in order to be fully prepared for the SAT. What test preparation companies are teaching is in fact genuine enrichment, albeit of a particular sort that is designed to fill the gap between the skills that students are taught and those that are tested on standardized tests. The better the test preparation program, the more it will address these gaps in curriculum alignment.

Overall Testing Strategies

The third element in good test preparation is instruction in test-specific overall testing strategies. For a given instrument, there should be instruction in how it is scored and in the most intelligent approaches to pacing, question selection and priority setting, guessing strategies, and so on, all in light of the particular structure of that test. Optimally, these strategies would be tailored to individual students' skills. Very different strategies will be warranted depending on the particularities of both the test and the student: a higher scoring student as opposed to a lower scoring one, a student taking the SAT as opposed to the ACT, a student strong in math as opposed to one strong in English, a student taking the test under normal time conditions as opposed to one taking it with special accommodations. Each of these factors will contribute to developing an optimal testing strategy for a particular student on a particular test.

Full-Length Practice Tests

The fourth element that adds to the effectiveness of a test preparation program is the administration of practice tests (ideally full-length ones) in the weeks prior to the administration of the actual test. Practice tests are useful for four reasons: to allow students (a) to put their overall testing skills into practice under timed conditions, (b) to learn to apply their newly learned problem-solving skills under pressure, (c) to desensitize themselves somewhat to the high-pressure environment of test day, and (d) to gain the stamina they need to perform with full concentration for the duration of the test. Let us look at each of these factors.

Practice of Generic and Test-Specific Testing Strategies

Once students have learned overall testing strategies: how to choose the ideal number of questions to attempt, how to prioritize and pace themselves through those questions, and so on, they then need an opportunity to put these strategies into practice under actual timed conditions in order to assimilate them fully. Many of these strategies revolve around pacing and timing considerations, so perfecting them requires actual timed practice. Likewise, students need to practice their particular problem-solving skills under actual test conditions, to ensure that they will be able to perform well when the real test day arrives.

Desensitization to the Stress of the Testing Environment

Test anxiety may cause students to fail to test to their ability, especially when they are faced with a high-stakes test (whether or not the test is actually used in such a fashion). Taking several tests under actual test conditions will help students acclimate to that sort of testing environment and make the actual test day less stressful.

Stamina

Although the sheer length of testing time required for many standardized tests probably does not elicit much sympathy, it probably should. Working with full concentration under extreme time pressure for hours on end is not a skill that students are given much room to develop. Students are typically accustomed to 45- or 55-minute examinations in their ordinary courses; the difference between those and a 180-minute (or longer) test is considerable. We should not

underestimate the degree to which physical training is an element of test preparation. Many students, whether due to fatigue or inability to concentrate, simply fade away in the final section or sections of a test.

This is why it is important that students take several full-length practice tests in the weeks leading up to the actual test—not wholly unlike training for a marathon or other feat of physical endurance. In the ideal case, they should take the practice test beginning at the same hour and under the same circumstances as the real test (in the case of the SAT, at 9 a.m. on a Saturday).

The qualifier *full-length* is important here. Many tests (such as the SAT) include an extra section that does not count toward a student's score, allowing the testing companies to pretest questions for future test administrations. This section is not identifiable, so students are forced to complete it with the same level of effort as all the other sections. Leaving aside questions about the ethicality of such a practice, it does have a significant effect on the length of the test: In the case of the SAT, it increases testing time from 2.5 to 3 hours. The practice tests prepared by testing companies (usually older versions of the test released in such books as *10 Real SATs* [College Board, 2000]) typically do not include this additional section and therefore do not represent a full-length test. From a purely physical standpoint, training for a 2.5-hour test is significantly different from training for a 3-hour test, which is 20 percent longer. Therefore providing sample tests that are not truly full length does the students a significant disservice in preparing them for test day; they may well end up fading out in the final 30 minutes of the test, which they have not adequately prepared for.

In all the programs I have studied, there has been a significant correlation between score improvement and the number of full-length practice tests administered during the program. The ETS and College Board studies also concur that repeated testing improves students' scores on their testing instruments. However, these studies usually assume that students retake the test months or years later, whereas during a test preparation program a student may take several tests within a four- to six-week period. I am aware of no evidence showing that one or the other method is any more effective in and of itself. I suspect, however, that repeated administrations during a four- to six-week period will facilitate students learning from their mistakes between one administration and the next. In any case, we can say with confidence that the more full-length practice tests offered in a test preparation program—up to four or five separate administrations—the more effective that program will be.

Instruction Oriented Toward the Specific Test

When students take a test such as the SAT that is not well aligned with their school curriculum, they often feel that they are poor students—that there was something they should have learned in school but did not—and that their poor score is therefore their own fault. But we have seen that many of the skills that apply to the SAT are not taught in school—especially SAT-specific testing strategies. Test preparation activities need to be labeled as such, so students understand that what they are now learning is something they were not supposed to have learned until now, and that these skills will have application primarily for the specific test they are to take.

Of course, the greater the misalignment between test and school curricula, the more severe this problem. When a given instrument is not aligned with a specific school curriculum, the test actually poses a social and educational policy problem for schools (and, indeed, for the whole educational establishment). To the degree that such tests are (often inappropriately) used to judge the performance of teachers, schools, and school districts, this creates an incentive for schools to begin doing more test-specific test preparation during the school day, which means diverting time from their standard curriculum. Bear in mind that the whole problem of "teaching to a test" arises *only* when the test is not curriculum aligned. When a test is well aligned with a school's curriculum (both in the concepts taught and in the logic of their presentation), then there is no gap between teaching to a test and teaching *tout court*. They are in that case one and the same.

Conclusion

I have presented the following five components as necessary to a good test preparation program. The more of these elements (and the higher their quality) in a program, the more likely that program will improve test scores:

1. Review of basic skills
2. Instruction in test-specific problem-solving skills
3. Instruction in overall test-taking strategies
4. Full-length practice tests
5. Instructional environment specifically oriented to testing

Does this mean that commercial test preparation is the only or the best option? Not necessarily. Students can get a number, if not all, of

these elements in various ways. They may get their basic skills from school and their test-specific problem-solving techniques and practice tests from a book or from an in-school program. Several in-school programs may have several but not all of these elements. The advantage of a professional service is that it delivers all of these elements in a high-quality way.

Of course, this chapter leaves many of the sticky ethical questions unanswered. Should test preparation be necessary? As a practical matter, it probably always will be: To the degree that any test is not quite curriculum aligned, there will be a place for relatively short-term test preparation to fill the gap between the students' learned curriculum and whatever the particular test instrument is testing. Does this make the test unfair? Only if test preparation is not considered an entitlement and not all students have equal access to it. In my personal opinion, this access should be secured equally, and at a national level, in the interests of public education: If we intend to test our students, we should do our best to ensure they have adequate preparation for these tests.

References

College Board. (2000). *10 real SATs.* New York: College Entrance Examination Board.

Evans, F. R. (1980). *A study of the relationships among speed and power aptitude test scores and ethnic identity* (Research Report RR-80–22). Princeton, NJ: Educational Testing Service.

Levine, A. E. (1979). *Effects of coaching on standardized admission examinations: Revised statistical analyses of data gathered by the Boston regional office of the Federal Trade Commission, Bureau of Consumer Protection.* Washington, DC: U.S. Government Printing Office.

Pike, L. W. (1978). *Short-term instruction, testwiseness, and the Scholastic Aptitude Test: A literature review with research recommendations.* New York: College Entrance Examination Board.

Further Reading

College Entrance Examination Board. (1965). *Effects of coaching on Scholastic Aptitude Test scores.* New York: College Entrance Examination Board.

Crouse, J., & Trusheim, D. (1988). *The case against the SAT.* Chicago: University of Chicago Press.

Owen, D. (1985). *None of the above: The myth of scholastic aptitude.* Boston: Houghton Mifflin.

Pike, L. W., & Evans, F. R. (1972). *Effects of special instruction for three kinds of mathematics aptitude items.* New York: College Entrance Examination Board.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

[ ] This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

[X] This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").